MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. $AD-A738525$	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On the Stability of Signal Detection		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A.F. Gualtierotti		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0491 N00014-81-K-0373
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of North Carolina Chapel Hill, NC 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-269 SRO 105
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics & Probability Program Office of Naval Research Arlington, VA 22217		12. REPORT DATE May 1983
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Signal detection; Stability; Contaminated Gaussian

document studies

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The probabilities of false alarm and false dismissal for sure signals embedded in additive noise, is studied. The noise is contaminated Gaussian, and the framework is such that continuous-time problems can be studied.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

84 02 21 053

ON THE STABILITY OF SIGNAL DETECTION

by

A. F. Gualtierotti*

Abstract

One studies probabilities of false alarm and false
dismissal for sure signals imbedded in additive noise
as functions of contamination when the noise is contami-
nated Gaussian.

*Department of Statistics, University of North Carolina, Chapel Hill, NC
27514.

IDHEAP, Sciences humaines 1, U. Lausanne, CH-1015 Lausanne.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	

By	
Distribution/	
Availability Codes	

Dist	Avail and/or Special
A-1	

83 11 28 094

1. INTRODUCTION

One of the basic problems in statistical communication theory is that of detecting, at the end of a communication channel, a signal imbedded in noise. The Gaussian model is a model for which the noise as well as the received signal are Gaussian processes with equivalent laws, so that the resulting detection problem is nonsingular. Unfortunately "most solutions to problems involving signals in Gaussian noise are based on a knowledge of the auto-correlation function or spectrum (if the noise is stationary). In fact, of course, as mentioned before, they are rarely, if ever, known precisely" (19). This observation motivates the paper just quoted and, in it, its author investigates the stability of the decision functions based on a model with Gaussian noise and covariance function R in two ways. First he considers the case of Gaussian noise with covariance function R+A instead of just R. However "the restriction to Gaussian measures were better removed even though the noise is nearly Gaussian, if the stability criteria are to be believable in practice" (19). Stability is then understood as a form of continuity of the decision functions when the topology is that of weak convergence, a concept which is now widely used in the field of robust statistical procedures (11). Of course "this formulation does not allow much hope for actually calculating useful bounds on errors" (19). Since the most important elements of a detection problem are the probabilities of false alarm and false dismissal, one may try to study the stability problem by taking an intermediate position between the two described above. This is the object of the present work and the idea is to look at specific contaminations of Gaussian laws in order to find analytic

expressions for the probability of false alarm and that of false dismissal as functions of contamination. The aim is not so much to produce models for specific Gaussian disturbances as to get some feel for what these can do to the probabilities of false alarm and false dismissal.

It shall be assumed that signals and noises have finite energy, that is are stochastic processes with square integrable paths. These will induce probability measures on the Borel sets of $L_2(0,1)$, the space of equivalence classes of square integrable functions. $L_2(0,1)$ is a standard separable Hilbert space and it is notationally simpler to work with a general separable Hilbert space H, which we shall do. $B(H)$ is the sigma algebra generated by the open sets of H and is the smallest sigma algebra which makes all continuous linear functionals $<\cdot,h>$, h in H, measurable. $<\cdot,\cdot>$ is the inner product of H and the corresponding norm is $||\cdot||$. All measures considered shall be defined on $B(H)$. They are typically obtained as follows. If $X(\omega,t)$ is a stochastic process defined on the probability space (Ω,A,μ) such that $\int_0^T X^2(\omega,t)\,dt$ is finite for μ-almost every ω, it induces, on $B(H)$, a measure P_X defined by

$$P_X\{x \text{ in } H: <x,h_1> \text{ in } B_1,\ldots,<x,h_n> \text{ in } B_n\} =$$
$$= P\{\omega \text{ in } \Omega: <X(\omega,\cdot),h_1> \text{ in } B_1,\ldots,<X(\omega,\cdot),h_n> \text{ in } B_n\},$$

where $h_1,\ldots,h_n$ belong to H and $B_1,\ldots,B_n$ belong to $B(\mathbb{R})$, the Borel sets of $\mathbb{R}$. Here $<X(\omega,\cdot),h> = \int_0^T X(\omega,t)\,h(t)\,dt$.

If S denotes the signal and N the noise, the detection problem $(N,S+N)$ is nonsingular if P_{S+N} and P_N are mutually absolutely continuous, that is $P_{S+N}(B) = 0$ if and only if $P_N(B) = 0$ for Borel B in $B(H)$. The singularity of a detection problem on a space of infinite dimension is notoriously unstable (19). Consequently,

for reasons analogous to those which require detection problems to be non-singular (3,14,18), attention shall be restricted to laws which are not only in the neighbourhood of Gaussian laws but also equivalent to these.

One tool which has proved useful in investigations dealing with robust statistical procedures is that of contamination (21). Let P be the law of the chosen model; its contaminations are of the form $P^C = (1-\varepsilon)P + \varepsilon Q$, where Q is some probability measure and $0 < \varepsilon < 1$. If $P = P_X$ and $Q = Q_Y$ and U is a random variable, independent of X and Y, which takes the value 1 with probability $1-\varepsilon$ and the value 0 with probability ε, then P^C is the law of $Z = UX + (1-U)Y$, that is the paths of Z are obtained by sampling among those of X and Y according to the law of U. In this paper we shall use two types of contaminations. The first type (T1-contaminations) is constructed with measures Q which are Gaussian and the second type (T2-contaminations) with measures Q which are absolutely continuous with respect to P and have a quadratic form as Radon-Nikodým derivative. The use of such simple contaminations has a number of advantages. First of all, their covariance has the form R+A so that we have a nongaussian generalization of the stability study within the Gaussian model, as in (19). Moreover, the situation is at the same time sufficiently simple so that one can obtain analytic information on the probabilities of interest, and sufficiently complex so that one can exhibit features which do not appear in the Gaussian perturbation model.

The first part of the paper deals with the basic properties of the contaminations just introduced. We show in particular that they share with Gaussian laws many properties, the main exception being that they do not necessarily

converge weakly to laws of the same type. In the second part of the paper we
use these contaminations to study the stability of the detection problem in case
of sure signal and additive Gaussian noise. Besides giving a number of quanti-
tative and qualitative results involving actual and nominal probabilities of
false alarm, we are able to show that weak convergence alone fails in general
to control the behaviour of the parameters which determine the probabilities of
false alarm: what really matters is the way in which this convergence takes place.
Finally, we show that the type of contamination which is considered plays an
important role. Thus contaminations by Gaussian laws result in worse patholo-
gies than contaminations of the second type considered: the former have in a
sense "too many degrees of freedom" for weak convergence to control.

We consider here only the detection of sure signals. For some results in the
case of stochastic signals, see (7). Proofs are outlined in the appendix.

2. CONTAMINATIONS

P will always denote a Gaussian probability on $B(H)$. It is determined by
its mean m_P and its covariance operator R_P, and we shall write $P \sim N(m_P, R_P)$.
m_P is an element in H and is uniquely determined by the relation

$$E_P <\cdot, h> = <m_P, h>, \text{ for all h in H.}$$

In case $P = P_X$, m_P is the mean function of X. The operator R_P is uniquely de-
termined by the relation

$$E_P \{<\cdot, h> - <m_P, h>\}\{<\cdot, k> - <m_P, k>\} = <R_P h, k>, \text{ for all h and k in H.}$$

In case $P = P_X$ and $R_X(s, t)$ is the covariance of X, R_P is defined by

$$R_P f)(t) = \int_0^T R_X(t, u) f(u) du.$$

Let $(a\otimes b)(x) = <b,x>a$. Then R_P has the representation

$$R_P = \sum_{k=1}^{\infty} \lambda_k \, e_k \otimes e_k,$$

where $\lambda_k > 0$, $\sum_{k=1}^{\infty} \lambda_k < \infty$ and $\{e_k\}$ is a complete orthonormal set in the closure

of the range of R_P. Since this latter set is the support of P, one can always sup-

pose it is H. e_k is an eigenvector of R_P corresponding to the eigenvalue λ_k. It

has been shown that Gaussian measures P on the Borel sets of $L_2(0,1)$ are all

of the form P_X, X a Gaussian process with paths in $L_2(0,1)$ (15). The Fourier

transform of P, FTP, is given by the relation

$$FTP(h) = E_P(\exp\{i<\cdot,h>\}) = \exp\{i<m_P,h> - \tfrac{1}{2}<R_P h, h>\}.$$

If $P = P_X$, FTP(th) is the characteristic function of the random variable $<X(\cdot,\cdot),h>$,

t real.

For T1-contaminations, we choose $Q \sim N(m_Q, R_Q)$, while for T2-contaminations

we choose $dQ = C^{-1} ||A^{\frac{1}{2}}(x-a)||^2 \, dP$. A is an operator which is linear, continuous,

positive and self-adjoint, a is a fixed element in H and $C = \text{trace}\{A(R_P + (m_P-a)\otimes(m_P-a))\}$.

ε is chosen to be $C(c+C)^{-1}$, where c is nonnegative. We have already indicated

that contaminations can be thought of as the laws of processes $Z = UX + (1-U)Y$,

where X has law P and Y has law Q. We are going to add the requirement that X

and Y have the same set of paths. One way to achieve this is to require that P

and Q be mutually absolutely continuous, so that they have the same support. In

this way, we discard cases such as those for which X would have continuous paths

and Y discontinuous ones, for which no doubt would exist as to which of P and

P^c obtains. For examples of T2-contaminations, see (8,9). In particular T2-con-

taminations of stationary processes yeld harmonizable ones. We thus have

Proposition 1

If P^c is a T1-contamination, P and P^c are mutually absolutely continuous if and only if P and Q are. P and its T2-contaminations are always mutually absolutely continuous./

The first assertion follows from the dichotomy theorem for Gaussian measures (10) and the second is a consequence of a zero-one law for Gaussian measures (2).

Assumption

P and P^c are henceforth assumed to be mutually absolutely continuous./

We shall first need the mean m_c and the covariance R_c of P^c. A direct calculation yields

Proposition 2

For T1-contaminations $\quad m_c = (1-\varepsilon)m_P + \varepsilon m_Q \quad$ and

$$R_c = (1-\varepsilon)R_P + \varepsilon R_Q + \varepsilon(1-\varepsilon)(m_P - m_Q) \otimes (m_P - m_Q),$$

and for T2-contaminations $m_c = m_P + 2(c+C)^{-1}R_P A(m_P - a)$ and

$$R_c = R_P + 2(c+C)^{-1}R_P A R_P - 4(c+C)^{-2}R_P A(m_P - a) \otimes (m_P - a)A R_P./$$

Remark 1

From Proposition 2 it follows that both types of contaminations have a mean and a covariance of the same form, that is

$m_c = m_P + R^{\frac{1}{2}}b$ and $R_c = R_P^{\frac{1}{2}}(I + T + \tau \, t \otimes t)R_P^{\frac{1}{2}}$.

(P and Q have been chosen equivalent in the T1 case so that in particular m_Q belongs to the range of $R_P^{\frac{1}{2}}$). These are essentially the conditions which ensure that a Gaussian law with mean m_c and covariance R_c is equivalent to P (16), so that, even if P_c is wrongly assumed to be Gaussian, there is no way to discriminate surely between P and P^c.

For every fixed h in H, one may define a real random variable $X\{h\}$ by setting $X\{h\}(x) = <x,h>$, x in H. For T1-contaminations, the law of $X\{h\}$ with respect to P^C, defined by

$$P^C_{X\{h\}}(B) = P^C(x \text{ in } H: X\{h\}(x) \text{ in } B), B \text{ in } B(\mathbb{R}),$$

is the obvious mixture of the laws of $X\{h\}$ with respect to P and Q. For T2-contaminations one has

Proposition 3

Let $\alpha(h) = 2(c+C)^{-1} <R_p A(m_p-a),h>$,

$\beta(h) = 2(c+C)^{-1} <R_p AR_p h,h>$,

$\phi(t) = \exp\{it<m_p,h>-\tfrac{1}{2}t^2<R_p h,h>\}$.

The characteristic function ϕ of the random variable $X\{h\}$ with respect to P^C, a T2-contamination, is given by

$$\phi(t) = \{1+i\alpha(h)t-\tfrac{1}{2}\beta(h)t^2\}\phi(t).$$

Let $a(h) = 1-(c+C)^{-1} \{<R_p AR_p h,h>/<R_p h,h>\}$,

$b(h) = 2(c+C)^{-1} \{<R_p A(m_p-a),h>/<R_p h,h>\}$,

$c(h) = (c+C)^{-1} \{<R_p AR_p h,h>/<R_p h,h>^2\}$,

$G(y)$ = Gaussian density with mean $<m_p,h>$ and variance $<R_p h,h>$.

The density g corresponding to ϕ is given by

$$g(y) = \{a(h)+b(h)(y-<m_p,h>)+c(h)(y-<m_p,h>)^2\}G(y). /$$

Remark 2

Suppose that $a = m_p$ and let $z = (y-<m_p,h>)/<R_p h,h>^{\frac{1}{2}}$,

$$L^2 = (c+C)^{-1} \{<R_p AR_p h,h>/<R_p h,h>\},$$

and $K^2 = \{2\pi<R_p h,h>\}^{\frac{1}{2}}$.

The density g of Proposition 3 can then be written

$$g(y) = \gamma(z) = K^2(1-L^2+L^2z^2)\exp\{-\tfrac{1}{2}z^2\}.$$

If $L^2 \leq 1/3$, for which a sufficient condition is $(c+C)^{-1}||R_P^{\frac{1}{2}}AR_P^{\frac{1}{2}}|| \leq 1/3$, γ has a unique maximum for $z = 0$. Otherwise γ presents two symmetric peaks. Laws with such densities have been shown to model some types of noise (17). Since $R_P^{\frac{1}{2}}AR_P^{\frac{1}{2}}$ is a compact operator, $<R_PAR_Ph,h>/<R_Ph,h> = <R_P^{\frac{1}{2}}AR_P^{\frac{1}{2}}k,k>$, with $k = R_P^{\frac{1}{2}}h/||R_P^{\frac{1}{2}}h||$, will be small in all but a finite number of directions. This shows the kind of departure from normality one can expect. A sink which would appear in the density of all functionals would eventually touch the origin, making the law singular with respect to the Gaussian one. To obtain two peaks with T1-contaminations, one would need distinct means. The peaks would then appear in every direction, but the sink would be controlled in size. Another interesting feature of the density g is that it is normal for h in the kernel of R_PAR_P. The measure P^C can thus exhibit a Gaussian behaviour in many directions without actually being Gaussian. This "Gaussianness" is controlled by the range of A. This feature is absent in the case of T1-contaminations.

One space of interest in linear inference problems is the closure in the real L_2-space $L_2(P^C)$ of the set $(X\{h\}, h$ in $H)$ of random variables. This subspace is called the linear space of P^C and is denoted $L(P^C)$. This terminology arises as follows: if $P^C = P_X$, linear operations on X are of the form $<X(\cdot,\cdot),h> = X\{h\}$. $L(P)$ contains only normal random variables. The next result states that $L(P^C)$ also contains random variables which all have a law of the same type.

Let P^C be a contamination and S_c be the operator $R_c+m_c\Theta m_c$. If f belongs to $L(P^C)$, the following relation defines an element of H denoted $h(f)$:

$$h_f = \int_H xf(x)P^C(dx).$$

For two such elements, $h(f)$ and $h(g)$, one can define the inner product

$$\langle h(f), h(g) \rangle_{H(P^C)} = E_{P^C}(fg).$$

Let $H(P^C) = \{h(f), f \text{ in } L(P^C)\}$ and define $U_{P^C} : L(P^C) \to H(P^C)$ by $U_{P^C}(f) = h(f)$.

$H(P^C)$ is then a Hilbert space, contained in H as a subset, and U_{P^C} is a unitary

operator. As a set, $H(P^C)$ is the range of $S_C^{\frac{1}{2}}$ and is isomorphic to the reprodu-

cing kernel Hilbert space of the process $\{X(h), h \text{ in } H\}$ (Lemma 1, Appendix).

Let now B be the operator from $L(P^C)$ to H defined by $Bf = R_P^{\frac{1}{2}} S_C^{-\frac{1}{2}} U_{P^C} f$. B is well

defined and continuous (Lemma 4, Appendix). Since $\langle R_C h, h \rangle \leq \langle S_C h, h \rangle$, $R_C^{\frac{1}{2}} = $

$S_C^{\frac{1}{2}} W_C$, where W_C is a bounded operator on H (5). Recalling that $m_C = m_P + R_P^{\frac{1}{2}} b$ and

that $\qquad R_C = R_P^{\frac{1}{2}}(I + T + t \cdot \otimes t) R_P^{\frac{1}{2}} = R_P^{\frac{1}{2}}(I + \hat{T}) R_P^{\frac{1}{2}}$ (Proposition 2 and Remark

1), one has

Proposition 4

Let f belong to $L(P^C)$ and h, h_n belong to H. Suppose that $U_{P^C} f = S_C^{\frac{1}{2}} h$ and

$h = \lim S_C^{\frac{1}{2}} h_n$ (We have assumed, which is no restriction, that P has full sup-

port. This implies that H is the closure of the range of $S_C^{\frac{1}{2}}$.). Then:

1) there exists m in H such that $\lim \langle m_P, h_n \rangle = \langle m, h \rangle$ (If m_P belongs to the

    range of $R_P^{\frac{1}{2}}$, $\langle m, h \rangle = \langle R_P^{-\frac{1}{2}} m_P, Bf \rangle$.);

2) $E_{P^C}(f) = \langle m, h \rangle + \langle b, Bf \rangle$ (If m_P belongs to the range of $R_P^{\frac{1}{2}}$, $E_{P^C}(f) = $

    $\langle b + R_P^{-\frac{1}{2}} m_P, Bf \rangle$.);

3) $V_{P^C}(f) = ||(I + \hat{T})^{\frac{1}{2}} Bf||^2 = ||W_C^* h||^2$;

4) For T1-contaminations, the characteristic function ϕ of f is given by

$\phi(t) = (1-\varepsilon)\phi_1(t) + \varepsilon\phi_2(t)$, where ϕ_1 is the characteristic function of a $N(<m,h>, ||Bf||^2)$-random variable and ϕ_2 is the characteristic function of a $N(<m,h> - <R_P^{-\frac{1}{2}}(m_P - m_Q), Bf>, ||W_c^* h||^2)$-random variable; for T2-contaminations, $\phi(t) = (1 + it<b, Bf> - \frac{1}{2}t^2 ||T^{\frac{1}{2}}Bf||^2)\phi_1(t)$. /

This result yields a useful computational tool, stated as a corollary.

Corollary

Let $f_1, \ldots, f_n$ belong to $L(P^C)$ and $h_1, \ldots, h_n$ solve $U_{P^C} f_i = S_c^{\frac{1}{2}} h_i$. Write $\vec{m}_1$, $\vec{m}_2$, $\vec{m}_3$ for the n-vectors with respective entries $<m, h_i>, -<B^* R_P^{-\frac{1}{2}}(m_P - m_Q), f_i>_{L(P^C)}$, $<B^* b, f_i>_{L(P^C)}$, and M_1, M_2, M_3 for the $(n \times n)$-matrices with respective entries

$<B^* Bf_i, f_j>_{L(P^C)}, <B^*(I + T_Q)Bf_i, f_j>_{L(P^C)}, <B^* TBf_i, f_j>_{L(P^C)}$, where T_Q solves

$R_Q = R_P^{\frac{1}{2}}(I + T_Q)R_P^{\frac{1}{2}}$. Then, if M_1 has an inverse, $(f_1, \ldots, f_n)^t$ has a density g given by, in case of T1-contaminations, $g(\vec{x}) = (1-\varepsilon)g_1(\vec{x}) + \varepsilon g_2(\vec{x})$, where g_1 is the density of a $N(\vec{m}_1, M_1)$-random vector and g_2 is the density of a $N(\vec{m}_1 - \vec{m}_2, M_2)$-random vector, and, in case of T2-contaminations,

$g(\vec{x}) = (1 - \text{trace}(M_1^{-1} M_2) - <\vec{x} - \vec{m}_1, M_1^{-1}\vec{m}_3> + \frac{1}{2}<M_1^{-1} M_3 M_1^{-1}(\vec{x} - \vec{m}_1), (\vec{x} - \vec{m}_1)>)g_1(\vec{x})$. /

There are two reasons which explain the effectiveness of Gaussian laws for solving problems of detection and estimation. The first is that their nonlinear space $L_2(P)$ is the direct sum of the symmetric tensor powers of their linear space $L(P)$ (12), which roughly means that the elements in $L_2(P)$ are limits of sums of polynomials of increasing order evaluated at an orthonormal basis for $L(P)$. The second reason is that the set of Gaussian laws is closed for the topology of weak convergence. It can be checked that the first result remains true

for T1- and T2- contaminations. The second however no longer obtains: the set
of contaminations is not closed for weak convergence, as the following results
show, and it is at the boundary that pathologies occur.

Proposition 5

Let $(P_n^C, n$ in $N)$ be a family of T1-contaminations which converge weakly to
the probability M. If lim inf $\varepsilon_n = 0$ or lim sup $\varepsilon_n = 1$, M is Gaussian. Otherwise
$M = (1-\mu)M_1 + \mu M_2$, where M_1 and M_2 are Gaussian and $0 < \mu < 1$. M is not neces-
sarily a T1-contamination and is Gaussian if and only if $M_1 = M_2$ (in particu-
lar M_1 and M_2 need not be mutually absolutely continuous)./

Proposition 6

Let $(P_n^C, n$ in $N)$ be a family of T2-contaminations which converge weakly to
the probability M. M has then a characteristic function ϕ given by

$$\phi(h) = (1 + i<d,h> - <S^{\frac{1}{2}}VS^{\frac{1}{2}}h,h>) \exp(i<n,h> - \tfrac{1}{2}<Sh,h>),$$

where S is a linear operator which is compact, nonnegative and selfadjoint and
V is linear and bounded (to have a T2-contamination, one needs an S which has
finite trace and a V of the form $S^{\frac{1}{2}}WS^{\frac{1}{2}}$). S has finite trace in one of the following
two cases:

1) $0 < \text{lim inf } \{C_n/(c_n+C_n)\} \leq \text{lim sup}\{C_n/(c_n+C_n)\} < 1;$

2) if $\alpha_n(h)$ is zero when $||R_n^{\frac{1}{2}}h||$ is and $(c_n+C_n)^{-1}(<R_n A_n R_n h,h>/<R_n h,h>)$
   otherwise and if $\beta_n(h)$ is zero when $||R_n^{\frac{1}{2}}h||$ is and $<b_n,h>/<R_n h,h>$ otherwise, th
   $\sup(\alpha_n(h),n,h) < 1$ and $\sup(|\beta_n(h)|,n,h) < 1./$

3. STABILITY OF SINGULARITY

Consider two Gaussian measures P_1 and P_2 representing respectively the law
of the noise and the law of the received signal in some detection problem. Detec-

tion is nonsingular if and only if $m_1 - m_2$ belongs to the range of the square root of $\frac{1}{2}(R_1 + R_2)$ and if $R_2 = R_1^{\frac{1}{2}}(I+T)R_1^{\frac{1}{2}}$ with T Hilbert-Schmidt and $\sigma(T) > -1$ ($\sigma(T)$ denotes the spectrum of T) (16). Suppose now the actual laws one deals with are P_1^C and P_2^C, T1- or T2-contaminations, possibly of a different type. The following result proves that the criterion for singularity which applies in the Gaussian case applies to T1- and T2-contaminations as well, so that singularity is stable over the class of T1- and T2-contaminations.

Proposition 7

Let P_1^C and P_2^C be T1- or T2-contaminations with respective means $m_{c,1}$ and $m_{c,2}$ and covariances $R_{c,1}$ and $R_{c,2}$. The detection problem (P_1^C, P_2^C) is either singular or nonsingular and it is singular if and only if $m_{c,1} - m_{c,2}$ belongs to the range of the square root of $\frac{1}{2}(R_{c,1} + R_{c,2})$ and if $R_{c,2} = R_{c,1}^{\frac{1}{2}}(I+S)R_{c,2}^{\frac{1}{2}}$, S Hilbert-Schmidt and $\sigma(S) > -1$. Furthermore, if the detection problem is nonsingular, $dP_2^C/dP_1^C = (dP_2^C/dP_2)(dP_2/dP_1)(dP_1/dP_1^C)$. /

4. DETECTION OF SURE SIGNALS

One observes a sure function which can have two possible sources: it is either a noise sample N or a distorted signal sample s+N, where s is known. The nominal model is that model for which the law P of the noise N is Gaussian with mean zero and covariance R_P of the form $\sum_n \lambda_n e_n \otimes e_n$, as defined in the introduction. The actual model is the model for which the law of N is P^C, with mean zero. The nominal model represents the model actually used and the actual model represents the system which is being investigated, the "true" model. The corresponding detection problem is nonsingular whenever s belongs to the range of the square root of R_P (6), which is isomorphic to the

reproducing kernel Hilbert space of R_P (Lemma 1, appendix). Since R_P and R_{P^C} have square roots with identical ranges(Lemma 4, appendix), the detection problem is nonsingular for the actual as well as for the nominal model. It shall henceforth be assumed that detection problems are nonsingular.

The maximum likelihood decision function for the nominal model is given by

$$D(X) = \sum_n (s_n/\lambda_n) \, X\{e_n\} \text{ with } s_n = <s,e_n> \text{ and } X\{e_n\}(y) = <y,e_n> \quad (19).$$

A. The case of T1-contaminations

By choice, $P^C = (1-\varepsilon)P + \varepsilon Q$ with $Q \sim N(m_Q, R_Q)$ and P and Q mutually absolutely continuous. Then $R_Q = R_P^{\frac{1}{2}}(I+T_Q)R_P^{\frac{1}{2}}$, T_Q Hilbert-Schmidt and $\sigma(T_Q) > -1$ (16). Since $m_{P^C} = m_c = 0$, $m_Q = 0$ (Proposition 2). Let $e_{n+1}, \ldots, e_{n+p}$ span the range of the projection $\pi_{n,p}$, $C_1 = ||\pi_{n,p} R_P^{-\frac{1}{2}}s||^2 + \varepsilon ||T_Q^{\frac{1}{2}}\pi_{n,p} R_P^{-\frac{1}{2}}s||^2$ and $C_2 = ||\pi_{n,p} R_P^{-\frac{1}{2}}s||^2$.

Then $E_{P^C}\{\sum_{i=1}^{n+p} (s_i/\lambda_i) X\{e_i\}\}^2$ equals C_1 if $X = N$ and equals C_1+C_2 if $X = s+N$, so that $D(X)$ belongs to $L_2(P^C)$. When $X = N$, let Y_1 be a normal random variable with mean zero and variance $||R_P^{-\frac{1}{2}}s||^2$, and, when $X = s+N$, let Y_1 have the same law and variance, but mean equal to the variance. Similarly, let Y_2 be normal with, when $X = N$, mean zero and variance $||R_Q^{-\frac{1}{2}}s||^2$, and, when $X = s+N$, mean $||R_P^{-\frac{1}{2}}s||^2$ and variance $||R_Q^{-\frac{1}{2}}s||^2$. Let f_1 and f_2 be the respective densities of Y_1 and Y_2. Then $D(X)$ has density $f(x) = (1-\varepsilon)f_1(x) + \varepsilon f_2(x)$. Finally let $p^a = P^C\{D(X) > \alpha | N\}$ and $p^n = P\{D(X) > \alpha | N\}$. Then $p^a - p^n = \varepsilon(Q\{D(X) > \alpha | X=N\} - p^n)$. Φ denotes the distribution function of a $N(0,1)$-random variable. The polar decomposition (22) yields $R_Q^{\frac{1}{2}} = R_P^{\frac{1}{2}}(I+T_Q)^{\frac{1}{2}}U$, where it is no restricition to suppose U unitary, so that $R_Q^{-\frac{1}{2}} = U^*(I+T_Q)^{-\frac{1}{2}}R_P^{-\frac{1}{2}}$. Let $\beta = \alpha/||R_P^{-\frac{1}{2}}s||^2$.

$\bar{s} = R_P^{-\frac{1}{2}} s / ||R_P^{-\frac{1}{2}} s||, \quad t = ||(I+T_Q)^{-\frac{1}{2}} \bar{s}||.$ Then $p^n - p^a = \epsilon \{ \phi(\beta) - \phi(\beta t) \}.$

Proposition 8

Fix s and $\beta > \phi^{-1}(3/4)$. Then

1) $\Delta = \sup \{ |p^a - p^n|, P^C \text{ a } T1\text{-contamination} \} = \phi(\beta) - \frac{1}{2};$

2) there exists $(P_k^C, k \text{ in } N)$ such that $W\text{-lim } P_k^C = P$ (W- indicating weak convergence) and $\lim_k |p_k^a - p_k^n| = \Delta$. /

The reason for 2) is that P and P^C can be close in the sense of weak convergence as well as close to being orthogonal. The result also snows that the discrepancy between p^a and p^n can increase with the level of the test.

Proposition 9

There exist sequences $\{P_k^C, s_k, k \text{ in } N\}$ such that $W\text{-lim } P_k^C = P$, $\lim s_k = 0$ and $\lim_k |p_k^a - p_k^n| = 1 - \phi(\beta)$. /

The reason for this phenomenon is that the class of admissible signals is too large for $|p^a - p^n|$ to be controlled by weak convergence.

Another way to compare p^a and p^n is to look at the ratio p^a/p^n. One has

$$p^a/p^n = 1 - \epsilon + \epsilon \{ (1 - \phi(\beta t)) / (1 - \phi(\beta)) \}.$$

As a function of t, p^a/p^n decreases from $1 - \epsilon + \frac{1}{2}(1 - \phi(\beta))^{-1}$ to $1 - \epsilon$. One thus has $p^a < p^n$ if $t > 1$ and $p^a > p^n$ if $t < 1$. When the level of the test increases, the behaviour of p^a ^{n} depends on the relative behaviour of β, ϵ and t. Suppose thus that t goes t   er than β goes to infinity: p^a/p^n is unbounded, so that p^n goes to ze   an p^a. If now ϵ goes to 1 and t goes to infinity, p^a/p^n goes to zero so that p^a goes to zero faster than p^n. We have seen in Propositions 8 and 9 that this behaviour can occur as P_k^C converges weakly to P.

As the examples considered show, weak convergence, when ϵ does not go

to zero, depends on the behaviour of $R_P^{\frac{1}{2}} T_Q R_P^{\frac{1}{2}}$ whereas the behaviour of $|p^a - p^n|$ and p^a/p^n depends on that of T_Q and these behaviours are not necessarily related.

Let now P_s^c be the law of $s+N$ and p^a be the probability of false alarm for the maximum likelihood detector for the actual detection model, that is $p^a = P^c\{dP_s^c/dP^c > \alpha\}$. Again one would like to know to what extent $p^n = P\{dP_s/dP > \alpha\}$ is a "good" approximation of p^a. Let $k = \{1-\epsilon+\epsilon(dQ/dP)\}^{-1}\{1-\epsilon+\epsilon(dQ_s/dQ)(dQ/dP)(dP/dP_s)\}$. Then $dP_s^c/dP^c = k(dP_s/dP)$, so that k represents the effect of contamination. The following assumptions allow one to obtain a manageable form for k and will be made for the remaining part of this section.

Assumptions

1) s belongs to the range of R_P (which is a subset of the range of $R_P^{\frac{1}{2}}$).

2) R_P and R_Q have the same range.

3) $R_Q^{-1} - R_P^{-1}$ has a symmetric closure G whose domain has P-probability one.

In particular, when $R_Q^{-1} - R_P^{-1}$ is bounded on the range of R_P, G is bounded and self-adjoint (16). Let $K(x) = (1-\epsilon)\det^{\frac{1}{2}}(I+T_Q)+\epsilon\exp\{-\frac{1}{2}<Gx,x>\}$. Then $k(x) = K(x-s)/K(x)$ (16). Thus, whenever s belongs to the kernel of G, the actual and the nominal probabilities are the same. One however has

Proposition 10

The range of k can be $\{x \text{ in } \mathbb{R}: 0 < x < \infty\}$. /

Proposition 10 shows that the values taken by the actual and nominal likelihoods can be quite different. Since the global behaviour of k, as a function of ϵ, G and T_Q is difficult to analyze, one may try to evaluate k at critical points. There are two of these of particular interest, 0 and s, at which one presumably would deci-

de that respectively no signal or a signal has been sent. Since $k(s) = 1/k(0)$, it suffices to consider $k(0)$. It is shown in the next proposition that $k(0)$ can be as large as $\exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$ which is the value taken by the nominal likelihood when the exact signal has been received. The actual likelihood has then the value one, whereas the nominal likelihood can be very small.

Proposition 11

$$\sup\{k(0)|\epsilon,Q\} \geq \exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}./$$

Here again, the bound can be approximated when P^c is close to P in the weak sense. Formulae for p^a are quite complicated and only in particular cases can one obtain global bounds for p^a in terms of p^n. Here is one case which shows what kinds of restrictions are then required.

Proposition 12

Let $\beta_k = 1+\{\epsilon_k/(1-\epsilon_k)\det^{\frac{1}{2}}(1+T_{Q_k})\}$. Suppose that G_k is nonnegative and that

$\beta_k \leq M < \infty$. Then, for $\alpha>0$ and $W\text{-lim } P_k^c = P$,

$$P\{dP_s/dP>\alpha M\} \leq \liminf P_k^c\{dP_{k,s}^c/dP_k^c>\alpha\} \leq \limsup P_k^c\{dP_{k,s}^c/dP_k^c>\alpha\} \leq P\{dP_s/dP>\alpha/M\}.$$

Thus a uniform bound on the admissible perturbations is required and the bounds obtained are good only if the perturbation is uniformly small.

B. The case of T2-contaminations

The features of T1-contaminations exhibited above indicate that stability results can only be obtained by restricting the class of signals and/or knowing some a priori uniform bounds on the perturbations one must consider (we have in particular assumed in Propositions 10, 11 and 12 - see the "Assumptions" preceding these statements - that the signal belongs to the range of R_p, rather than to the range of $R_p^{\frac{1}{2}}$, which is imposed by a nonsingular problem). In this sense, T1-contamina-

tions have "too many degrees of freedom". One can thus attempt to obtain stability

results by restricting the class of contaminations considered. This is what one a-

chieves by introducing T2-contaminations. Whereas ϵ and Q act as independent

variables for T1-contaminations, for T2-contaminations, ϵ depends on Q in the sense

that $\epsilon = C/(c+C)$, where C is a function of the parameters which determine Q (see

in particular the proof of Proposition 13). Intuitively, T2-contaminations are ob-

tained from T1-contaminations by keeping the first terms of the Taylor expansion

of the likelihood of Q with respect to P. This indicates that classes of contaminations

intermediate between T1- and T2-contaminations can be obtained by keeping more

terms in the Taylor expansion considered above.

It is again assumed that $m_c = 0$, which is achieved by setting $R_p Aa = 0$. As in

the case of T1-contaminations, one may check that $D(X)$ belongs to $L_2(P^C)$. Let Y_1

be the same random variable as in A. and f_1 be its density. Let $\epsilon = ||A^{\frac{1}{2}}s||^2/(c+C)||R_p^{-\frac{1}{2}}$

and $f_2(x) = x^2 f_1(x)/||R_p^{-\frac{1}{2}}s||^2$. With respect to P^c, $D(X)$ has then a density f given

by the relation $f(x) = (1-\epsilon)f_1(x)+\epsilon f_2(x)$. Let I_A be the indicator of the closure of

the range of A and B, p^a, p^n and $\bar{s}$ have the same meaning as in A. One then has,

if $K = ||A^{\frac{1}{2}}R_p^{\frac{1}{2}}\bar{s}||^2/\{c+\text{trace}(AR_p)+||A^{\frac{1}{2}}a||^2\}$, $p^a-p^n = K I_A(s)\int_B^\infty (x^2-1)\exp(-\frac{1}{2}x^2)dx/\sqrt{2\pi}$.

The next proposition shows that the pathologies illustrated in Proposition 8 and 9

cannot occur in the case of T2-contaminations. The reason is clear: ϵ is not inde-

pendent of Q. There is a restriction however: s is required to belong to the range

of R_p.

Proposition 13

Let s belong to the range of R_p and $W\text{-lim } P_k^C = P$, then $\lim p_k^a = p^n$./

A simple limit calculation (4) gives the relative behaviour of p^a and p^n: p^n tends

to zero faster than p^a.

Proposition 14

For fixed ϵ and large a, $p^a/p^n \sim a^2$./

Finally, we consider again the actual likelihood and the actual probability of false alarm, still denoted p^a. In contrast to the case of T1-contaminations, the function k of Proposition 10 is bounded for T2-contaminations. Indeed, if $K(x) = c+!!A^{\frac{1}{2}}(x-a)!$ $dP^C = \{K \quad /(c+C)\}dP$, so that, if $k(x) = K(x-s)/K(x)$, $dP^C_s/dP^C = k(dP_s/dP)$. Then

Proposition 15

Let a_A belong to the kernel of A, $\lambda_m = \frac{1}{2}\{1+(1+4c/<As,s>)^{\frac{1}{2}}\}$ and $\lambda_M = \frac{1}{2}\{1-(1+4c/<As,s>)\}$ Then $k(a+\lambda_m s+a_A) \leq k(x) \leq k(a+\lambda_M s+a_A)$./

From Proposition 15 one can get a result which is more satisfactory than Proposition 12, but which is still insufficient in the sense that it depends on a priori knowledge on the perturbations, which is unlikely to be available.

Proposition 16

If s is in the range of R_p and $\lim \sup C_n/c_n < \infty$, then, when $W\text{-}\lim P^C_n = P$, $\lim P^C_n\{dP^C_{n,s}/dP^C_n>\alpha\} = P\{dP_s/dP>\alpha\}$./

5. SUMMARY

Let $p_1 = P\{dP_s/dP>\alpha\}$, the probability of false alarm for the Gaussian model, $p_2 = P^C\{dP_s/dP>\alpha\}$, the probability of false alarm for the "exact" model when the detector is the likelihood for the Gaussian model, and $p_3 = P^C\{dP^C_s/dP^C>\alpha\}$, the probability of false alarm for the "exact" model. We have considered the case of a Gaussian detection model which is only an approximation to an "exact" contaminated model and compared the quantities p_1, p_2, and p_3 when the law P^C of the "exact" model is close in the topology of weak convergence to P, the law of the

nominal model, which in the present case is Gaussian. We have seen that the quantities $|p_1-p_i|$ and p_1/p_i, $i=2,3$, depend on the type of contamination as well as on the distance between P and P^c. Proposition 9 shows for example that one can have, for a T1-contamination, $|p_1-p_2| \geq \epsilon > 0$, though s and the distance between P and P^c are negligible. Proposition 16 shows however that under fairly mild restrictions, $|p_1-p_3|$ is small as soon as P and P^c are close and P^c is a T2-contamination. We have also shown that looking at the likelihood can be rather misleading when the model is not "exact".

6. OUTLINE OF PROOFS

The proof of Proposition 4 is based on four lemmas which we first state.

Lemma 1

Let S_c, $H(P^c)$ and U_{p^c} be defined as in Proposition 4. For f in $L(P^c)$ define

$F_f: H \to \mathbb{R}$ by the relation $F_f(h) = \int_H f(x)X\{h\}(x)P^c(dx)$. Let $H(S_c) = \{F_f, f \text{ in } L(P^c)\}$, $\langle F_f, F_g \rangle_{H(S_c)} = E_{p^c}(fg)$ and $U_{S_c}: L(P^c) \to H(S_c)$ be defined by the relation $U_{S_c}f = F_f$.

Then

1) $H(S_c)$ is a Hilbert space (the reproducing kernel Hilbert space of X) and U_{S_c} is a unitary operator,

2) $||h(f)|| \leq ||S_c^{\frac{1}{2}}||^{\frac{1}{2}}||h(f)||_{H(P^c)}$,

3) the two sets $H(P^c)$ and range$(S_c^{\frac{1}{2}})$ are equal,

4) $\langle S_c^{\frac{1}{2}}h, S_c^{\frac{1}{2}}k \rangle_{H(P^c)} = \langle h,k \rangle$, so that $S_c^{\frac{1}{2}}: H \to H(P^c)$ is unitary,

5) if $U_{p^c}f = S_c^{\frac{1}{2}}h$ and $f_n = X\{h_n\}$, then $f = \lim f_n$ if and only if $h = \lim S_c^{\frac{1}{2}}h_n$, in $L(P^c)$ and H respectively./

Lemma 1 follows from the definitions and first principles. These also yield

Lemma 2

If R_P is injective, $\text{trace}(AR_P) = ||R_P^{\frac{1}{2}}AR_P^{\frac{1}{2}}||$ if and only if $A = h \otimes h$./

Lemma 3

If P^C is a T2-contamination and $T + \tau\, t \otimes t$ is the operator of Proposition 2,

$\sigma(T + \tau\, t \otimes t) > -1$./

pf. : It suffices to prove that -1 is not an eigenvalue of $T + \tau\, t \otimes t$ and this follows from

the inequality $1 + \{(1 + <T, h\ h>)^{\frac{1}{2}} - 1\} \leq ||b|| \leq 1$./

Lemma 4

Let $\hat{T} = T + \tau\, t \otimes t$. Then

1) $R_C^{\frac{1}{2}}$ and $R_P^{\frac{1}{2}}$ have the same range and $R_C^{\frac{1}{2}} = R_P^{\frac{1}{2}}(I + \hat{T})^{\frac{1}{2}}V_C$, where V_C is unitary.

2) the map $B: L(P^C) \rightarrow H$ defined by the relation $Bf = R_P^{\frac{1}{2}}S_C^{-1}U_{P^C}f$ is linear and continuous./

pf. : 1) follows from Lemma 3, (5) and the polar decomposition (22). B is well

defined on H^* and, if f is in H^*, $||Bf|| \leq K||f||_{L(P^C)}$./

Proposition 4

pf.: Let $f_n = X\{h_n\}$. Then $f = \lim f_n$ by Lemma 1, so that, by Lemma 4, $Bf = \lim R_P^{\frac{1}{2}}h_n$.

Furthermore, $E_{P^C}(f) = \lim <m_P, h_n> + <b, Bf>$, so that $\lim <m_P, h_n>$ is a continuous

linear functional on $L(P^C)$. The first part of 1) follws by the Riesz theorem and

Lemma 1. If m_P belongs to the range of $R_P^{\frac{1}{2}}$, $<m_P, h_n>$ tends to $<R_P^{-\frac{1}{2}}m_P, Bf>$. The result

on the variance follows from Lemma 4 and that on the characteristic function from

Proposition 3./

Proposition 5

pf. : For $\epsilon, x,$ and y in the open unit interval, let $f_\epsilon(x, y) = (1 - \epsilon)x + \epsilon y$. If $0 < \alpha < 1$,

there is a β in the open unit interval such that $f_\epsilon(x, y) \geq 1 - \beta$ implies $x \geq 1 - \alpha$ and

$y \geq 1-\alpha$. Consequently, if $\{P_\lambda^C, \lambda \text{ in } \Lambda\}$ is relatively compact, so are $\{P_\lambda, \lambda \text{ in } \Lambda\}$

and $\{Q_\lambda, \lambda \text{ in } \Lambda\}$, provided $\varepsilon_\lambda = \varepsilon$ and ε is in the open unit interval.

Suppose now $\{P_\lambda^C, \lambda \text{ in } \Lambda\}$ converges weakly to P^C so that $0 < \lim \inf \varepsilon_\lambda < \lim \sup \varepsilon_\lambda < 1$.

If $\{P_n, n \text{ in } N\}$ is a subsequence of $\{P_\lambda, \lambda \text{ in } \Lambda\}$, let $\{\varepsilon_{n_k}\}$ converge to ε in the

open unit interval. Set $P_{n_k}^{C, \varepsilon} = (1-\varepsilon)P_{n_k} + \varepsilon Q_{n_k}$. $\{P_{n_k}^{C, \varepsilon}\}$ converges weakly to P^C.

The remarks at the beginning imply that $\{P_{n_k}\}$ is relatively compact. The proposi-

tion follows. To see that the limit is not necessarily a contamination, consider $\{P_n^C\}$

with $P_n^C = (1-\varepsilon)P + \varepsilon Q_n$, $P \sim N(0, R_P)$, $Q_n \sim N(0, R_P^{\frac{1}{2}}(I+T_n)R_P^{\frac{1}{2}})$, $T_n = ((1/n)-1)u \otimes u$,

u a unit vector. /

Proposition 6

<u>pf</u>. : Because of (9), P^C has characteristic function given by the relation

$\phi(h) = \{1 + i<b, h> - <Bh, h>\} \exp\{i<m, h> - \frac{1}{2}<Sh, h>\}$, where B and S are bounded

linear operators. Since the convolutions $\{P_n^C * P_n^C\}$ converge weakly to $\{P^C * P^C\}$,

the latter has characteristic function $|\phi(h)|^2 = \{<b, h>^2 + (1-<Bh, h>)^2\} \exp\{-<Sh, h>\}$.

$|\phi(h)|^2$ is continuous in the S-topology and thus, if $\{x_n\}$ converges weakly to

zero, $\lim |\phi(x_n)|^2 = 1$. Consequently, $\lim \sup <Sx_n, x_n> < \infty$. If this limit is

positive, choosing if necessary a subsequence, one may assume that $\lim <Sx_n, x_n> = \sigma^2 > 0$.

One then has that $\lim \{1 - t^2 <Bx_n, x_n>\} = \exp\{\sigma^2 t^2\}$, so that $\lim <Bx_n, x_n> = 0$ and

consequently that $\lim <Sx_n, x_n> = 0$. $S^{\frac{1}{2}}$ is thus compact and thus so is S. From the

inverse Fourier transform of ϕ, one obtains that $<Bh, h> \leq <Sh, h>$, so that

$B = S^{\frac{1}{2}} T S^{\frac{1}{2}}$, T bounded.

1) yields, as in the proof of Proposition 5, that $\{P_n\}$ is relatively compact, so

that $\{R_n\}$ is compact. 2) yields that P_n^C has a Fourier transform ϕ_n such that

$|\phi(h)|^2 \leq \exp\{-(1-\delta)<R_n h, h>\}$, which also implies that $\{R_n\}$ is compact. That

S has finite trace follows then from (20)./

Proposition 7

pf. : Recall that P and P^c have been assumed equivalent. Using the result (1) "λ and μ orthogonal and λ and ν equivalent imply μ and ν orthogonal", one can see that the problem (P_1^c, P_2^c) is either singular or nonsingular and that it is nonsingular if and only if the problem (P_1, P_2) is nonsingular.

Suppose then that the problem (P_1, P_2) is nonsingular. The operators R_1, R_2, $\frac{1}{2}(R_1+R_2)$ have all a square root with the same range (16) and, by Lemma 4 of 6., $R_{c,i}$ and $R_i, i=1,2$, have a square root with the same range. Thus $m_{c,1} - m_{c,2} = m_1 - m_2 + R_1^{\frac{1}{2}} b_1 - R_2^{\frac{1}{2}} b_2$ belongs to the range of the square root of $\frac{1}{2}(R_{c,1} + R_{c,2})$. The polar decomposition and the equality $R_2 = R_1^{\frac{1}{2}}(I+T)R_1^{\frac{1}{2}}$ yield $R_2^{\frac{1}{2}} = R_1^{\frac{1}{2}}(I+T)^{\frac{1}{2}}V$, where V is unitary ($R_1$ may be assumed injective by restricting attention to the support of P_1). Using again Lemma 4 of 6., one can write $R_{c,2} = R_{c,1}^{\frac{1}{2}} W R_{c,1}^{\frac{1}{2}}$, where W is a product of operators which are either unitary or of the form $(I+U)^{\pm\frac{1}{2}}$, with U Hilbert-Schmidt and $\sigma(U) > -1$. Unitary operators always appear in pairs involving the operator and its adjoint. Since $(I+U)^{\pm\frac{1}{2}}$ can always be written in the form $I+\tilde{U}$ with $\tilde{U}$ Hilbert-Schmidt, W has the form $I+S$, S Hilbert-Schmidt. To see that $\sigma(S) > -1$, it is sufficient to remember that $R_{c,1}$ and $R_{c,2}$ have square roots with identical ranges, so that $I+S$ is invertible. The reverse implication is proved similarly./

Proposition 8

pf. : Let $H(\bar{s})$ be the subspace spanned by $\bar{s}$ and $H(\bar{s})^+$ be its orthogonal subspace. Let $\bar{\pi}(\bar{s})$ denote the projection onto $H(\bar{s})$ and $\pi(\bar{s})^+$ the projection onto $H(\bar{s})^+$. Set $T_Q = \theta\pi(\bar{s})$. Then $t = (1+\theta)^{\frac{1}{2}}$. Since θ can be any real number strictly

larger than -1, t can take any positive value. As a function of t, $\varepsilon|\Phi(\beta)-\Phi(\beta t)|$ is a convex function with a minimum at $t = 1$ and extreme values at $t = 0$, where its value is $\varepsilon\{\Phi(\beta)-\frac{1}{2}\}$, and at $t = \infty$, where its value is $\varepsilon\{1-\Phi(\beta)\}$. Furtermore, for $\Phi(\beta) > 3/4$, $\Phi(\beta)-\frac{1}{2} > 1-\Phi(\beta)$.

$\{P_k^c\}$ converges weakly to P if and only if either $\{\varepsilon_k\}$ converges to zero or $\{Q_k\}$ converges weakly to P. If the former obtains, $\lim |p_k^a - p_k^n| = 0$. Let $\{\alpha_k\}$ be a sequence of positive numbers whose sum is finite. Define $\theta_{k,l} = 0$ for $l = 1,\ldots,k-1$, $\theta_{k,k} = (1/\theta_k)-1$, θ_k being defined below, and, for $l > k$, $\theta_{k,l} = \alpha_k$. Let $\bar{s}_k = \langle\bar{s},e_k\rangle$ and choose $\{\theta_k\}$ such that $\lim \theta_k \bar{s}_k^2 = \infty$. Then $\sum_l \{\bar{s}_k^2/(1+\theta_{k,l})\} \geq \theta_k \bar{s}_k^2$, so that $t_k \leq \{\theta_k \bar{s}_k^2\}^{-\frac{1}{2}}$ and $\lim t_k = 0$. Now, $\langle R_{P_k^c}h,h\rangle = \langle R_P h,h\rangle + \varepsilon_k \sum_{l=k}^{\infty} \theta_{k,l}\lambda_l \langle e_l,h\rangle^2$.

Since the sequence $\{\theta_{k,l}\}$ is uniformly bounded, $\lim \langle R_{P_k^c}h,h\rangle = \langle R_P h,h\rangle$.

Furthermore trace$(R_{P_k^c})$ is bounded and $\sum_{l=m}^{\infty} \langle R_{P_k^c}e_l,e_l\rangle$ converges uniformly to zero as m increases to infinity. This insures that W-$\lim P_k^c = P$ (13) ./

Proposition 9

pf. : Set $s_k = \lambda_k^{\frac{1}{2}}e_k$. Then $||R_P^{-\frac{1}{2}}s_k|| = 1$, $\beta = \alpha$ and $\bar{s}_k = R_P^{-\frac{1}{2}}s_k/||R_P^{-\frac{1}{2}}s_k|| = e_k$. Let $T_{Q_k} = \theta_k \Pi(\bar{s}_k)$. Then, as in the proof of Proposition 8, $t_k = (1+\theta_k)^{\frac{1}{2}}$, $\langle R_{P_k^c}h,h\rangle = \langle R_P h,h\rangle + \varepsilon_k \theta_k \lambda_k$, trace$(R_{P_k^c})$ = trace $(R_P)+\varepsilon_k \theta_k \lambda_k$, and $\sum_{l=m}^{\infty} \langle R_{P_k^c}e_l,e_l\rangle \leq \sum_{l=m}^{\infty} \langle R_P e_l,e_l\rangle+\varepsilon_k \theta_k \lambda_k$. It thus suffices to choose $\{\theta_k\}$ such that $\lim \theta_k = \infty$ and $\lim \theta_k \lambda_k = 0$. /

Proposition 10

pf. : Fix g so that it is not orthogonal to s and $\Gamma > -||R_P^{\frac{1}{2}}g||^2$. Set $\bar{g} = R_P^{\frac{1}{2}}g/||R_P^{\frac{1}{2}}g||$.

and $\gamma = -\Gamma||R_p^{\frac{1}{2}}g||^2/(1+\Gamma||R_p^{\frac{1}{2}}g||^2)$. Let $\pi(\bar{g})$ be the projection onto $H(\bar{g})$ and

$\pi(\bar{g})^\perp$ be that onto $H(\bar{g})^\perp$. Then, if $T_Q = \gamma\pi(\bar{g})$, $(I+T_Q)^{-1} = (1+\Gamma||R_p^{\frac{1}{2}}g||^2)\pi(\bar{g})+\pi(\bar{g})^\perp =$

$I+R_p^{\frac{1}{2}}(\Gamma g\otimes g)R_p^{\frac{1}{2}}$. So $G = \Gamma g\otimes g$ and $K(x) = (1-\varepsilon)+\varepsilon(1+\Gamma||R_p^{\frac{1}{2}}g||^2)^{\frac{1}{2}}\exp\{-\frac{1}{2}\Gamma<g,x>^2\}$.

Set $<g,x> = y$, $<g,s> = a$, $-\frac{1}{2}\Gamma y^2 = P(y)$. k can then be expressed in the form

$k^*(y) = \{A+B\exp\{P(y-a)\}\}/\{A+B\exp\{P(y)\}\}$. If $L(y) = P'(y)\exp\{P(y)\}/\{A+B\exp\{P(y)\}\}$,

then $(d/dy)k^*(y) = Bk^*(y)\{L(y-a)-L(y)\}$. Now, for $\Gamma<0$, L is strictly increasing,

so that the graphs of $L(y-a)$ and of $L(y)$ are parallel and do not meet (because of

the assumption that a is different from zero, that is, g and s are not orthogonal).

The derivative of k^* has thus constant sign and k^* is a monotone function which

is unbounded and positive./

Proposition 11

pf. : It suffices to establish that $\sup\{k(0)|\varepsilon,G\} = \exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$. One has

that $<Gs,s> = ||R_p^{-\frac{1}{2}}s||^2<R_p^{\frac{1}{2}}GR_p^{\frac{1}{2}}\bar{s},\bar{s}>$. Since $(I+T_Q)^{-1} = I+R_p^{\frac{1}{2}}GR_p^{\frac{1}{2}}$, $<R_p^{\frac{1}{2}}GR_p^{\frac{1}{2}}\bar{s},\bar{s}> > -1$

and thus $-\frac{1}{2}<Gs,s> \leq \frac{1}{2}||R_p^{-\frac{1}{2}}s||^2$. Consequently

$k(0) \leq \{\varepsilon+(1-\varepsilon)\det^{\frac{1}{2}}(I+T_Q)\}^{-1}\{(1-\varepsilon)\det^{\frac{1}{2}}(I+T_Q)+\varepsilon\exp\{-\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}\}$. For a and b

positive and c strictly larger than one, the function $(a+bcx)/(a+bx)$ is increasing

and bounded by c. Thus $k(0) \leq \exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$. On the other hand, if

$\lambda = (1-\varepsilon)\det^{\frac{1}{2}}(I+T_Q)$ and $\mu = \varepsilon$,

$k(0) = (\lambda+\mu)^{-1}\{\lambda\exp\{-\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}+\mu\exp\{-\frac{1}{2}<(I+T_Q)^{-1}\bar{s},\bar{s}>\}\}\exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$.

So, by convexity,

$k(0) \geq \exp\{-\frac{1}{2}(\lambda+\mu)^{-1}\{\lambda||R_p^{-\frac{1}{2}}s||^2+\mu||(I+T_Q)^{-\frac{1}{2}}\bar{s}||^2\}\}\exp\{\frac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$.

If H_n is the subspace spanned by $e_1,\ldots,e_n$, π_n the projection onto H_n and $\pi_n^\perp$

that onto $H_n^\perp$, set $T_Q = n\pi_n$ and $\varepsilon_n = 1-\{1/(1+n)\}^n$. Then $\det^{\frac{1}{2}}(I+T_Q) = (1+n)^{\frac{1}{2}n}$

and $||(I+T_Q)^{-\frac{1}{2}}\bar{s}||^2 = (1+n)^{-1}||\pi_n\bar{s}||^2+||\pi_n^\perp\bar{s}||^2$. Taking the limit as n grows

yields $k(0) \geq \exp\{\tfrac{1}{2}||R_p^{-\frac{1}{2}}s||^2\}$./

Proposition 12

<u>pf.</u> : When C_n is nonnegative, $\beta_n^{-1} \leq k_n(x) \leq \beta_n$ and thus

$P_n^C\{dP_s/dP > \alpha M\} \leq P_n^C\{dP_{n,s}^C/dP_n^C > \alpha\} \leq P_n^C\{dP_s/dP > \alpha/M\}$. Since dP_s/dP is a

continuous function, the result follows from the properties of weak convergence.

Proposition 13

<u>pf.</u> : Let $h^* = R_p^{\frac{1}{2}}h/||R_p^{\frac{1}{2}}h||$. One then has that $E_Q\{\exp\{iX(h)\}\} =$

$\{1 - <R_p^{\frac{1}{2}}AR_p^{\frac{1}{2}}h^*, h^*><R_ph, h>/\{trace(AR_p) + ||A^{\frac{1}{2}}a||^2\}\}\exp\{-\frac{1}{2}<R_ph, h>\}$. So, if

$W\text{-}\lim Q_n = P$ and k is a vector in the range of $R_p^{\frac{1}{2}}$, with norm equal to one,

$\lim \varepsilon_n = \lim ||A_n^{\frac{1}{2}}R_p^{\frac{1}{2}}k||^2/\{trace(A_nR_p) + ||A_n^{\frac{1}{2}}a_n||^2\} = 0$./

Proposition 14

<u>pf.</u> : To obtain the extreme values of k, one studies the function $k(s^*+tx)$ as a

function of t. It turns out that s^* must be of the form $s^* = a + \lambda s + a_A$, with a_A in

the kernel of A, for s^* to be extremal. λ is obtained using this representation./

Proposition 15

<u>pf.</u> : If λ_m corresponds to s^*_{min} and λ_M to s^*_{max} in the proof of Proposition 14,

then $P^C\{D(X) > \log\alpha - \log k(a + \lambda_m s + a_A)\} \leq P^C\{dP_s^C/dP^C > \alpha\} \leq P^C\{D(X) > \log\alpha - \log k(a + \lambda_M s + a_A)\}$.

Then, if $d = 4c/<As, s>$, $k(a + \lambda_m s + a_A) = \{1 + (1+d)^{\frac{1}{2}}\}/\{(1+d)^{\frac{1}{2}} - 1\}$, which increases

strictly from zero to one as d goes from zero to infinity. If now $W\text{-}\lim P_n^C = P$

and $\lim \sup C_n/c_n < \infty$, $\lim <A_n s, s>/4c_n = 0$ and $\lim k(a_n + \lambda_{n,m} s + a_{A_n}) = 1$./

7. REFERENCES

(1) R.B. Ash, Real Analysis and Probability. New York, NY: Academic Press, 1972.

(2) C.R. Baker, "Zero-one laws for Gaussian measures on Banach space," Trans.

   Amer. Math. Soc. 186, pp. 291-308, 1973.

(3) C.R. Baker, "Conditions for absolute continuity," in Probability on Banach Spaces II, Lecture Notes in Math. no. 709. New York, NY: Springer Verlag, 1980.

(4) J. Dieudonné, Calcul Infinitésimal, Paris, France: Hermann, 1968.

(5) R.C. Douglas, "On majorization, factorization, and range inclusion of operators in Hilbert space," Proc. Amer. Math. Soc. 17, pp. 413-415, 1966.

(6) U. Grenander, "Stochastic processes and statistical inference," Ark. Mat. 1, pp. 195-277, 1950.

(7) A.F. Gualtierotti, "Sur la détection des signaux, la quantité d'information, et la capacité de retransmission, quand le bruit n'est pas exactement Gaussien," C.R. Acad. Sc. Paris 288A, pp. 69-71, 1979.

(8) A.F. Gualtierotti, "On Shannon information, and channel capacity for a class of contaminated Gaussian noises," Information and Control 46, pp. 46-70. 1980.

(9) A.F. Gualtierotti, "A class of stochastic processes with a law generalizing a nondecomposable law on the real line," Stochastic Process. Appl. 13, pp. 87-117, 1982.

(10) J. Hajék, "On a property of normal distributions of any stochastic process," Czech. Math. J. 8, pp. 610-618, 1958.

(11) P. Huber, Robust Statistical Procedures. CBMS-NSF Reg. Conf. Ser. Appl. Math. 27, Philadelphia, PA: SIAM, 1977.

(12) J. Neveu, Processus Aléatoires Gaussiens. Montreal, Canada: Les Presses de l'Université de Montréal, 1968.

(13) K.R. Parthasarathy, Probability Measures on Metric Spaces. New York, NY: Academic Press, 1967.

(14) E. Parzen, "Probability density functionals and reproducing kernel Hilbert

spaces," Proc. Sympos. Time Series Analysis. New York, NY: Wiley, 1963.

(15) B.S. Rajput and S. Cambanis, "Gaussian processes and Gaussian measures,"
Ann. Math. Stat. 45, pp,1944-1954, 1972.

(16) C.R. Rao and V.S. Varadarajan, "Discrimination of Gaussian processes,"
Sankhya A 25, pp. 303-330, 1963.

(17) G.A. Richards, "Simulation of signal and noise in a nonlinear channel,"
NATO Advanced Study Institute: Communication systems and random process
theory. Darlington, England: 1977.

(18) W.L. Root, "Singular Gaussian measures in detection theory," Proc. Sympos.
Time Series Analysis, New York, NY: Wiley, 1963.

(19) W.L. Root, "Stability in signal detection problems," Proc. Sympos. Appl. Math.
XVI. Providence, RI: Amer. Math. Soc. 1964.

(20) B. Simon, Trace Ideals and their Applications. Cambridge, England: Cambridge
Univ. Press, 1979.

(21) J.W. Tukey, "A survey of sampling from contaminated distributions,"
Contributions to Probability and Statistics. Stanford, CA: Stanford Univ.
Press, 1960.

(22) J. Weidmann, Linear Operators in Hilbert Spaces. New York, NY: Springer
Verlag, 1980.